

ORIGINAL ARTICLE

## Reasons for ordering laboratory tests and relationship with frequency of abnormal results

PAUL H. H. HOUBEN<sup>1</sup>, RON A. G. WINKENS<sup>1,2</sup>, TRUDY VAN DER WEIJDEN<sup>1</sup>, RENEE C. R. M. VOSSEN<sup>3</sup>, ANDRÉ J. M. NAUS<sup>4</sup> & RICHARD P. T. M. GROL<sup>1,5</sup>

<sup>1</sup>Maastricht University, School of Public Health and Primary Care (CAPHRI), Department of General Practice, <sup>2</sup>Maastricht University Medical Centre, Integrated Care Unit, <sup>3</sup>Maasland Hospital Sittard, Department of Clinical Chemistry & Haematology, <sup>4</sup>Laurentius Hospital Roermond, Department of Clinical Chemistry & Haematology, and <sup>5</sup>Radboud University Nijmegen, Scientific Institute for Quality of Healthcare

### Abstract

**Objective.** Laboratory tests are ordered on a daily basis, even though disease probability is often very low. Abnormal results, especially mildly abnormal results, can be difficult to interpret in these circumstances. Further insights into the occurrence of abnormalities can help improve rational test ordering and test interpretation. The objective was therefore to examine the frequency of mildly and markedly abnormal results and their relationship with physicians' reasons for ordering tests. **Design.** Prospective study. **Participants.** A total of 87 primary care physicians in the Netherlands collected data on 1775 patients. **Main outcome measures.** The physicians recorded the reason for ordering the tests, the most probable diagnosis and the pretest probability. The laboratories' reference values and specified "action limits" were used to assess the number of abnormal results and markedly abnormal results, respectively. **Results.** Laboratory results were received for 1621 patients and 15,603 tests were reported (mean 9.6). The proportion of abnormal test results increased with increasing pretest probability (from 13.9% to 34.7%) and was 13.4% for tests ordered to reassure the patient and 13.3% for psychosocial diagnoses. The proportion of patients with at least one abnormal test result was high: 53.1% for tests ordered to reassure and 57.7% in patients with low pretest probability. Corresponding values for a marked abnormality were 11.1% and 12.4%, respectively. **Conclusion.** Abnormal laboratory test results were frequent, even when pretest probability was low. Physicians should therefore carefully consider when tests are necessary. Future research could explore physicians' interpretation of test results and its impact on diagnosis and management.

**Key Words:** *Clinical chemistry tests, diagnostic tests, family practice, interpretation of laboratory tests, laboratory utilization, primary healthcare*

In primary care, correct interpretation of abnormal test results may be difficult. For instance, abnormal results can sometimes lead to an unjustified cascade of further investigations [1,2], while on the other hand they may erroneously fail to be further investigated [3].

These difficulties with the interpretation of abnormal results may have several causes. For example, the pretest probability of disease is usually low in primary care. Also, physicians may suspect that complaints and symptoms are related to psychosocial problems [4,5]. Many tests are indeed

ordered for reasons such as excluding pathology or reassuring the patient [6–8]. Furthermore, many abnormal lab results represent only minor deviations, and physicians may doubt their clinical relevance [9]. It is in particular the combination of minor abnormalities and a low pretest probability which may be difficult to interpret, as it is often easier to decide on the basis of more markedly abnormal results.

While several authors reported high percentages of abnormal laboratory results [10,11], the relationship with the physicians' reasons for ordering tests

Tests are often ordered even though disease probability is low. The frequency of abnormal results in these situations was unknown.

- Even when tested for reasons such as “patient reassurance”, many patients have abnormal results.
- Most abnormalities are only slightly abnormal.
- Physicians should carefully consider when tests are necessary. Future research could explore the interpretation of results and the impact on diagnosis and management.

and the frequency of mildly or markedly abnormal results are unknown. More information on this relationship could be important for education as well as to improve rational test ordering. It may also help in exploring areas of potential difficulty in the interpretation of results.

The aim of this study was therefore to examine the frequency of mildly and markedly abnormal results in routine practice, and to investigate how they relate to the reasons for ordering laboratory tests.

## Material and methods

### Design and setting

We conducted a prospective study among 87 primary care physicians and their patients in seven rural, suburban, and urban areas in the south of the Netherlands. Physicians' age, sex, and working time were representative of those in the Netherlands according to data from the Netherlands Institute

for Health Services Research (NIVEL). To prevent selection bias, each participating physician was personally instructed to include and to record data on the first 25 consecutive adult patients for whom they had decided to order laboratory tests. Also, to prevent physicians from changing their ordering patterns, it was stressed that we did not intend to measure performance or to give feedback on test ordering. Physicians working part time included a smaller number of patients, proportional to the number of hours a week they were working. Patients were asked to give informed consent.

### Measurements

The physicians recorded data when they ordered the laboratory tests, using forms that were specifically designed for the study and took about two minutes to complete. The forms had been pilot-tested and evaluated as regards validity, reliability, and user convenience in an iterative process among a sample of 10 primary care physicians and a questionnaire expert. The laboratories provided the researchers with a copy of the laboratory results.

### Outcome variables

**Test results.** We used the reference values from the laboratories, adjusted for age and sex where appropriate, to determine the number of abnormal results. We asked an expert group to help us discriminate between minor and marked abnormalities. This expert group comprised two clinical chemists, one primary care physician with expertise in laboratory testing, and two primary care physicians from the research team. The experts went through a consensus procedure to specify action limits for tests,

Table I. Reference values, action limits, and abnormal results.

Test	N tests	% abnormal test results according to reference values (95% CI)	Action limits	% markedly abnormal test results according to action limits (95% CI)
ALAT (U/l)	440	10% (8%–13%)	<70	3% (1%–4%)
Cholesterol (mmol/l)	584	38% (34%–41%)	<8.0	3% (1%–4%)
Creatinine clearance (ml/min/1.73 m <sup>2</sup> ) <sup>1</sup>	780	61% (58%–65%)	<60	21% (18%–24%)
CRP (mg/l)	175	38% (31%–45%)	<50	5% (1%–8%)
ESR (mm/hr)	877	22% (19%–25%)	<40	5% (3%–6%)
Ferritin (µg/l)	57	47% (34%–60%)	15–1000	35% (23%–47%)
Gamma GT (U/l)	239	24% (19%–30%)	<90	6% (3%–9%)
Glucose (mmol/l)	949	8% (6%–9%)	3.0–11.0	2% (1%–3%)
Hemoglobin (mmol/l)	878	13% (11%–16%)	♀ 6.5–10.5, ♂ 7.5–11.5	2% (1%–3%)
Leukocytes (*10 <sup>9</sup> /l)	574	13% (10%–25%)	2.5–15.0	2% (1%–3%)
MCV (fl)	716	6% (5%–8%)	75–105	1% (0%–2%)
Potassium (mmol/l)	334	16% (12%–20%)	3.0–5.5	1% (0%–3%)
Sodium (mmol/l)	137	7% (2%–11%)	130–150	0%
TSH (mU/l)	660	11% (9%–13%)	0.1–8.0	5% (3%–7%)

Note: <sup>1</sup>MDRD formula, adjusted to the laboratory method for creatinine measurement (jaffe/enzymatic).

Table II. Distribution of abnormal test results at test level.<sup>1</sup>

	N tests	% abnormal test results according to reference values	% markedly abnormal tests according to action limits
Patients (n = 1484)	7400	20.7%	4.9%
Sex			
Men	2857	21.4%	4.3%
Women	4543	20.3%	5.3%
Age			
< 40 years	1747	13.7%	2.5%
40–60 years	2773	19.8%	4.0%
> 60 years	2880	25.8%	7.4%
Reason for ordering tests			
Confirm/determine treatment	725	27.2%	9.1%
Exclude/physician's uncertainty	3654	18.6%	4.2%
Reassure/patient's request	1139	13.4%	2.0%
Monitoring (screening/check-up)	1761	27.0%	6.6%
Other reasons	33	15.2%	3.0%
Estimated pretest probability <sup>2</sup>			
Definitely no disease	541	13.9%	2.2%
Probably no disease	1990	14.2%	2.6%
Possibly disease	1651	19.9%	4.5%
Probably disease	1127	24.3%	7.0%
Definitely disease	196	34.7%	14.3%
Presumed diagnosis <sup>2</sup>			
A97, P, Z	1083	13.3%	2.7%
Other ICPC codes	4350	20.1%	4.9%

Notes: <sup>1</sup>Limited to the 14 tests in Table I. <sup>2</sup>The physicians did not estimate a pretest probability or mention a presumed diagnosis if they ordered the tests for screening or check-up.

indicating values for which they considered urgent medical intervention to be necessary. The action limits were established in an iterative process and were based as much as possible on evidence-based guidelines and the literature. After three discussion rounds, consensus regarding the action limits for 14 tests was reached (see Table I).

#### Independent variables

*Reason for ordering tests.* We distinguished nine reasons for test ordering, which were chosen on the basis of a qualitative interview study among primary care physicians [12]. Physicians recorded the most important reason for ordering the investigations by ticking one of nine boxes. We summarized these into five categories: (1) to exclude disease and reduce the physician's uncertainty, (2) to confirm diagnosis and to determine treatment, (3) to reassure patients and at patient's request, (4) monitoring of disease (screening for hypertension/cholesterol/diabetes and check-up for known disorder), and (5) other reasons.

*Estimated pretest disease probability.* The form physicians had to complete at the time of test ordering asked: "Do you suspect a disease?" The physicians answered on a five-point Likert scale: "definitely not", "probably not", "maybe", "probably yes" and "definitely yes".

*Presumed diagnosis.* Physicians recorded their most likely diagnosis at the time of test ordering. We coded these according to the International Classification of Primary Care (ICPC), defining two groups: (1) the most likely diagnosis refers to non-somatic disease, i.e. ICPC chapters P (psychological) and Z (social) and ICPC code A97 ("no disease"); and (2) all other diagnoses.

#### Analysis

We used SPSS 15 to analyze the data. We counted the number of abnormal test results according to the reference values and the number of markedly abnormal test results based on the action limits (test level). We also counted the number of patients with at least one abnormal or one markedly abnormal test result (patient level). We used the Chi-square test to test for significant differences in the numbers of patients with abnormal results within the categories of age, sex, reason for ordering tests, presumed diagnosis, and estimated pretest probability.

#### Results

A total of 87 primary care physicians participated in the study, and they included 1775 patients. We

received no laboratory results for 154 patients (8.7%), the most important reasons being failure to visit the laboratory (31.8%) and failure to give informed consent (13.6%), while the reason could not be retrieved in 42.9% of the cases. The laboratories reported 15,603 tests for the remaining 1621 patients. The mean of reported tests was 9.9 for tests ordered to exclude disease, 8.9 for tests ordered to reassure, and 7.9 for tests ordered for monitoring. There were 2875 (18.4%) abnormal test results according to the reference values. The 1621 patients included 440 (27.1%) with one abnormal test result, 310 (19.1%) with two abnormal test results, and 437 (27.0%) with three or more abnormal test results.

The tests for which action limits had been specified accounted for 47.4% ( $n=7400$ ) of all reported tests ( $n=15,603$ ). There were 1484 ( $1484/1621=91.5\%$ ) patients for whom at least one of these tests was ordered. (Table I) shows how often each test was reported and the percentage of abnormal and markedly abnormal test results.

At test level, the frequency of markedly abnormal test results was clearly smaller than the frequency of abnormal test results (Table II). The frequency of abnormal test results increased with increasing age (13.7%–25.8%) and increasing pretest probability (13.9%–34.7%), and was high for tests ordered to confirm disease (27.2%) and for screening (27.0%).

At patient level, the frequency of patients with one or more abnormal test results was considerable (Table III). Even if the physician had ordered to reassure, 53.1% of the patients had abnormal results and 11.1% had a markedly abnormal result. Patients' age and sex, the reason for ordering tests, and the estimated pretest probability were significantly related to the percentage of patients with abnormal results.

## Discussion

This study shows that many patients have abnormal laboratory test results, even when the tests are ordered in a situation of low pretest probability, such as to reassure patients (53.1%) or to exclude disease (62.5%). A large proportion of the abnormal test results were only mildly abnormal. Given the low pretest probability and the statistical definition of reference values, there is a fair chance that the abnormalities may not have clinical significance in terms of diagnosis or therapy [13]. This underlines the necessity for physicians to carefully consider their orders for laboratory tests. It also raises the question of what exactly physicians do with these results. Future research should therefore attempt to explore how physicians interpret these test results, and what the diagnostic and therapeutic consequences are.

Table III. Distribution of abnormal result at patient level.<sup>1</sup>

Reasons for ordering tests	N patients	% patients with abnormal results according to reference values (95% CI)	% patients with markedly abnormal results according to action limits (95% CI)
Patients	1484	64% (61%–66%)	21% (19%–23%)
Sex			
Men	591 (40%)	63% (60%–67%)	18% (15%–21%)*
Women	893 (60%)	64% (61%–67%)	23% (20%–26%)
Age			
< 40 years	343 (23%)	48% (42%–53%)*	10% (7%–13%)*
40–60 years	582 (39%)	60% (57%–64%)	16% (13%–20%)
> 60 years	559 (38%)	77% (74%–81%)	33% (29%–37%)
Reason for ordering tests	16 missing		
Confirm/determine treatment	167 (11%)	68% (61%–75%)*	30% (23%–37%)*
Exclude/physician's	651 (44%)	63% (59%–66%)	20% (17%–23%)
Uncertainty	207 (14%)	53% (46%–60%)	11% (7%–15%)
Reassure/patient's request	434 (29%)	69% (65%–73%)	24% (20%–27%)
Monitoring (screening/check-up)	9 (1%)	33% (7%–69%)	11% (0%–48%)
Other reasons			
Estimated pretest probability <sup>2</sup>	28 missing		
Definitely no disease	97 (9%)	58% (48%–68%)*	12% (6%–19%)*
Probably no disease	353 (34%)	54% (48%–59%)	14% (10%–17%)
Possibly disease	301 (29%)	66% (60%–71%)	22% (17%–27%)
Probably disease	218 (21%)	69% (63%–75%)	28% (22%–33%)
Definitely disease	53 (5%)	70% (57%–82%)	40% (26%–53%)
Presumed diagnosis <sup>2</sup>	38 missing		
A97, P, Z	173 (17%)	58% (51%–66%)	16% (10%–21%)
Other ICPC codes	839 (83%)	62% (59%–65%)	21% (18%–24%)

Notes: \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . <sup>1</sup>Limited to the 14 tests in Table I. <sup>2</sup>The physicians did not estimate a pretest probability or mention a presumed diagnosis if they ordered the tests for screening or check-up.

The total proportion of patients with an abnormal test result in our study was 73.2%. On the whole, our results are in line with those of others in terms of the high frequency of abnormal results found [10,11,14,15]. However, the test for ferritin was abnormal for comparatively large numbers of patients. This may be due to the fact that many laboratories in the Netherlands only report ferritin when hemoglobin is too low. The low number of abnormal test results when action limits are applied may relate to the fact that in these cases the statistical chance of finding an abnormality is very limited.

Strong points of this study were that it included many physicians and that they were a representative sample of those in the Netherlands. Also, the data were prospectively collected and we tried to prevent selection bias by careful instruction of the physicians. However, Dutch primary care physicians tend to order fewer laboratory tests than those in other countries [16]. In countries where physicians tend to order more tests there will probably be even more patients with abnormal results, underlining the need for rational test ordering and a better understanding of the influence of results on further diagnosis and therapy.

Since the action limits that were specified by the expert group might be criticized, we also examined the possibility of defining action limits by using 99% reference values, critical differences and a computerized expert system (Valab) [17]. However, none of these methods allowed us to specify action limits for our data corresponding to clinically significant abnormalities. Furthermore, varying our action limits caused only slight changes in the number of markedly abnormal results, which would not have changed our conclusions. We therefore think that the use of the action limits defined by the expert group was the best option for this study.

The widespread test ordering in the presence of low pretest probability may be discussed. It usually adds little information [18] and it increases costs while many test orders in low pretest probability circumstances may be considered unnecessary [1,2,19–26]. Also, some evidence indicates that it may give rise to diagnostic problems due to the frequent occurrence of marginally abnormal results [1,2,5]. However, at the same time, sporadic but important laboratory abnormalities may be found, despite the low pretest probability. So, reducing test ordering in low pretest probability may reduce costs, but it must be done carefully to avoid under-diagnosis.

A way to overcome the problem of misleading test results may be by using decision limits instead of reference values. Decision limits are based on consensus linking test results to clinical outcomes. A problem is that with different analytical systems used

by laboratories the decision limits may also vary, just as reference values may vary. Also, decision limits vary for specific clinical cases. This is complex to present and may negatively influence the presentation of laboratory reports. In fact, there is still much discussion in clinical laboratory literature, and in other areas such as diagnostic imaging, with regard to appropriate helpful methods for interpretation of test results, as all methods meet specific difficulties [27,28]. In their training, physicians often learn to interpret results by means of computations involving pretest probability, sensitivity, and specificity, which allow the posttest probability to be calculated [29,30]. This may, however, be difficult to apply in the complex environment of routine practice, where the results of several tests are usually reported, and where several test results may be abnormal [31]. Greater insight into the principles that underlie the interpretation of results in routine care might help to develop better methods to support physicians in interpreting abnormal results.

In conclusion, physicians are likely to be confronted with many marginally abnormal test results, even if tests were ordered to reassure or in the case of low pretest probability. Given the low pretest probability and the statistical definition of reference values, there is a fair chance that these abnormalities have no clinical significance. Therefore, physicians should consider carefully whether ordering a test is necessary. Furthermore, because the consequences of these abnormalities are unclear with respect to diagnosis and further management, future research should help to better understand physicians' interpretation of test results and its impact on diagnosis and management.

## Ethics

The Medical Ethics Committee of Maastricht University approved the study. There was no external funding. There are no conflicts of interest.

## References

- [1] Vafiadis P. The dilemma of the unexpected result. *Aust Fam Physician* 1996;25:971–3.
- [2] Deyo RA. Cascade effects of medical technology. *Annu Rev Public Health* 2002;23:23–44. Epub 2001 Oct 25.
- [3] Sherwood P, Lyburn I, Brown S, Ryder S. How are abnormal results for liver function tests dealt with in primary care? Audit of yield and impact. *BMJ* 2001;322:276–8.
- [4] Little P, Cantrell T, Roberts L, Chapman J, Langridge J, Pickering R. Why do GPs perform investigations? The medical and social agendas in arranging back X-rays. *Fam Pract* 1998;15:264–5.
- [5] Aiarzaguena JM, Grandes G, Salazar A, Gaminde I, Sanchez A. The diagnostic challenges presented by patients

- with medically unexplained symptoms in general practice. *Scand J Prim Health Care* 2008;26:99–105.
- [6] Van der Weijden T, van Bokhoven MA, Dinant GJ, van Hasselt CM, Grol RP. Understanding laboratory testing in diagnostic uncertainty: A qualitative study in general practice. *Br J Gen Pract* 2002;52:974–80.
  - [7] Van Bokhoven MA, Pleunis-van Empel MC, Koch H, Grol RP, Dinant GJ, van der Weijden T. Why do patients want to have their blood tested? A qualitative study of patient expectations in general practice. *BMC Fam Pract* 2006;13:7–75.
  - [8] Van Boven K, Dijksterhuis P, Lamberts H. Defensive testing in Dutch family practice: Is the grass greener on the other side of the ocean? *J Fam Pract* 1997;44:468–72.
  - [9] Patt CH, Yoo HY, Dibadij K, Flynn J, Thuluvath PJ. Prevalence of transaminase abnormalities in asymptomatic, healthy subjects participating in an executive health-screening program. *Dig Dis Sci* 2003;48:797–801.
  - [10] Mold JW, Aspy CB, Lawler FH. Outcomes of an insurance company-sponsored multichannel chemistry screening initiative. *J Fam Pract* 1998;47:1110–7.
  - [11] Phillips WR, Thompson DJ. Multi-channel laboratory testing and the unexpected abnormal result: A statistical myth corrected. *N Z Med J* 1981;94:462–4.
  - [12] Houben PHH, van der Weijden T, van Bokhoven MA, Droog A, Winkens RA, Grol R. Overwegingen van huisartsen bij het interpreteren van laboratoriumonderzoek; een kwalitatief onderzoek [How general practitioners interpret laboratory test results; a qualitative interview study]. *English summary. Huisarts Wet* 2005;48:736–32.
  - [13] Brigden ML, Heathcote JC. Problems in interpreting laboratory tests: What do unexpected results mean? *Postgrad Med* 2000;107:145–6, 51–2, 55–8. *passim*.
  - [14] Hubbell FA, Frye EB, Akin BV, Rucker L. Routine admission laboratory testing for general medical patients. *Med Care* 1988;26:619–30.
  - [15] Rutten GEHM, De GrauwWJC, Nijpels G, Goudswaard AN, Uitewaal PJM, Van der Does FEE, et al. NHG-standaard Diabetes Mellitus type II. *Huisarts Wet* 2006;49:137–52.
  - [16] Leurquin P, Van Casteren V, De Maeseneer J. Use of blood tests in general practice: A collaborative study in eight European countries. Eurosentinel Study Group. *Br J Gen Pract* 1995;45:21–5.
  - [17] Franzini C. Relevance of analytical and biological variations to quality and interpretation of test results: Examples of application to haematology. *Ann Ist Super Sanita* 1995; 31:9–13.
  - [18] De Ruijter W, Assendelft WJ, Macfarlane PW, Westendorp RG, Gussekloo J. The additional value of routine electrocardiograms in cardiovascular risk management of older people. *Scand J Prim Health Care* 2008;26:147–53.
  - [19] Van Wijk M, Mosseveld M, van der Lei J. Design of a decision support system for test ordering in general practice: Choices and decisions to make. *Methods Inf Med* 1999; 38:355–61.
  - [20] Smith BJ, McNeely MD. The influence of an expert system for test ordering and interpretation on laboratory investigations. *Clin Chem* 1999;45:1168–75.
  - [21] Bugter Maessen AM, Winkens RA, Grol RP, Knottnerus JA, Kester AD, Beusmans GH, et al. Factors predicting differences among general practitioners in test ordering behaviour and in the response to feedback on test requests. *Fam Pract* 1996;13:254–8.
  - [22] Verstappen WH, van der Weijden T, Sijbrandij J, Smeele I, Hermesen J, Grimshaw J, et al. Effect of a practice-based strategy on test ordering performance of primary care physicians: A randomized trial. *JAMA* 2003;289:2407–12.
  - [23] Lyon AW, Greenway DC, Hindmarsh JT. A strategy to promote rational clinical chemistry test utilization. *Am J Clin Pathol* 1995;103:718–24.
  - [24] Winkens RA, Pop P, Bugter Maessen AM, Grol RP, Kester AD, Beusmans GH, et al. Randomised controlled trial of routine individual feedback to improve rationality and reduce numbers of test requests. *Lancet* 1995;345:498–502.
  - [25] Van Walraven C, Naylor CD. Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits [see comments]. *JAMA* 1998;280:550–8.
  - [26] Zaat JO, Schellevis FG, van Eijk JT, van der Velden K. Do out-of-office laboratory tests affect diagnoses in general practice? *Scand J Prim Health Care* 1995;13:46–51.
  - [27] Henny J. Interpretation of laboratory results: The reference values, a necessary evil? *Clin Chem Lab Med* 2007;45: 939–41.
  - [28] Espeland A, Baerheim A. General practitioners' views on radiology reports of plain radiography for back pain. *Scand J Prim Health Care* 2007;25:15–9.
  - [29] Puhan MA, Steurer J, Bachmann LM, ter Riet G. A randomized trial of ways to describe test accuracy: The effect on physicians' post-test probability estimates. *Ann Intern Med* 2005;143:184–9.
  - [30] Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychol Rev* 1995;102:684–704.
  - [31] Reid MC, Lane DA, Feinstein AR. Academic calculations versus clinical judgments: Practicing physicians' use of quantitative measures of test accuracy. *Am J Med* 1998;104: 374–80.